Title (en)
Discriminative training of models for sequence classification

Title (de)
Unterschiedliche Ausbildung von Modellen für die Folgeklassifizierung

Title (fr)
Formation discriminative de modèles pour classification de séquence

Publication
EP 1939758 A2 20080702 (EN)

Application
EP 07122900 A 20071211

Priority
US 64698306 A 20061228

Abstract (en)
Classification of sequences, such as the translation of natural language sentences, is carried out using an independence assumption. The independence assumption is an assumption that the probability of a correct translation of a source sentence word into a particular target sentence word is independent of the translation of other words in the sentence. Although this assumption is not a correct one, a high level of word translation accuracy is nonetheless achieved. In particular, discriminative training is used to develop models for each target vocabulary word based on a set of features of the corresponding source word in training sentences, with at least one of those features relating to the context of the source word. Each model comprises a weight vector for the corresponding target vocabulary word. The weights comprising the vectors are associated with respective ones of the features; each weight is a measure of the extent to which the presence of that feature for the source word makes it more probable that the target word in question is the correct one.

IPC 8 full level
G06F 17/28 (2006.01)

CPC (source: EP US)
G06F 40/44 (2020.01 - EP US)

Citation (applicant)
- F. OCH; H. NEY: "Discriminative training and maximum entropy models for statistical machine translation", PROCEEDINGS OF ACL, 2002
- K. YAMADA; K. KNIGHT: "A syntax-based statistical translation model", PROCEEDINGS OF 39TH ACL, 2001
- P. BROWN ET AL.: "The Mathematics of Machine Translation: Parameter Estimation", COMPUTATIONAL LINGUISTICS, vol. 16, no. 2, 1993, pages 263 - 312
- F.J. OCH; H. NEY: "A systematic comparison of various statistical alignment models", COMPUTATIONAL LINGUISTICS, vol. 29, no. 1, 2003, pages 19 - 51
- S. KUMAR; W. BYRNE: "A weighted finite state transducer implementation of the alignment template model for statistical machine translation", PROCEEDINGS OF HLT-NAACL 2003, May 2003 (2003-05-01)
- P. KOEHN; F. J. OCH; D. MARCU: "Statistical phrase-based translation", PROCEEDINGS OF THE HUMAN LANGUAGE TECHNOLOGY CONFERENCE 2003, May 2003 (2003-05-01)
- N. BERTOLDI ET AL.: "The ITC-IRST Statistical Machine Translation System for IWSLT-2004", PROCEEDINGS OF THE INTERNATIONAL WORKSHOP ON SPOKEN LANGUAGE TRANSLATION (IWSLT), September 2004 (2004-09-01), pages 51 - 58
- R. ZENS ET AL.: "The RWTH Phrase-based Statistical Machine Translation System", PROCEEDINGS OF THE INTERNATIONAL WORKSHOP ON SPOKEN LANGUAGE TRANSLATION (IWSLT), October 2005 (2005-10-01), pages 155 - 162
- S. BANGALORE; G. RICCARDI: "Stochastic finite-state models for spoken language machine translation", MACHINE TRANSLATION, vol. 17, no. 3, 2002
- F. CASACUBERTA; E. VIDAL: "Machine translation with inferred stochastic finite-state transducers", COMPUTATIONAL LINGUISTICS, vol. 30, no. 2, 2004, pages 205 - 225
- S. KANTHAK; H. NEY: "Fsa: An efficient and flexible c++ toolkit for finite state automata using on-demand computation", PROCEEDINGS OF THE 42ND ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS, 2004, pages 510 - 517
- J. M. CREGO; J. B. MARINO; A. DE GISPERT: "An ngram-based statistical machine translation decoder", PROC. OF THE 9TH EUROPEAN CONF. ON SPEECH COMMUNICATION AND TECHNOLOGY (INTERSPEECH'05), September 2005 (2005-09-01), pages 3185 - 3188
- F. J. OCH; H. NEY: "Discriminative training and maximum entropy models for statistical machine translation", PROC. ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS, July 2002 (2002-07-01), pages 295 - 302, XP002524733
- V. GOFFIN ET AL.: "The AT&T WATSON Speech Recognizer", PROCEEDINGS OF ICASSP, 2005
- A. STOLCKE: "SRILM - An Extensible Language Modeling Toolkit", PROC. INTL. CONF. SPOKEN LANGUAGE PROCESSING, 2002
- J. LAFFERTY; A. MCCALLUM; F. PEREIRA: "Conditional random fields: Probabilistic models for segmenting and labeling sequence data", PROCEEDINGS OF ICML, 2001
- A.L. BERGER ET AL.: "A Maximum Entropy Approach to Natural Language Processing", COMPUTATIONAL LINGUISTICS, vol. 22, no. 1, 1996, pages 39 - 71
- R. MALOUF: "A comparison of algorithms for maximum entropy parameter estimation", PROCEEDINGS OF CONLL-2002, 2002, pages 49 - 55
- M. DUDIK; S. PHILLIPS; R.E. SCHAPIRE: "Proceedings of COLT'04", 2004, SPRINGER VERLAG, article "Performance Guarantees for Regularized Maximum Entropy Density Estimation"
- A. MCCALLUM; D. FREITAG; F. PEREIRA: "Maximum entropy Markov models for information extraction and segmentation", PROC. 17TH INTERNATIONAL CONF. ON MACHINE LEARNING, 2000, pages 591 - 598
- L. BOTTOU: "Une Approche théorique de l'Apprentissage Connexionniste: Applications a la Reconnaissance de la Parole", PH.D. THESIS, 1991
- J. LAFFERTY; A. MCCALLUM; F. PEREIRA: "Conditional random fields: Probabilistic models for segmenting and labeling sequence data", PROC. 18TH INTERNATIONAL CONF. ON MACHINE LEARNING, 2001, pages 282 - 289
- J. PLATT: "NIPS", 1999, MIT PRESS, article "Probabilistic outputs for support vector machines and comparison to regularized likelihood methods"

Cited by
CN104484322A

Designated contracting state (EPC)
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU LV MC MT NL PL PT RO SE SI SK TR

Designated extension state (EPC)
AL BA HR MK RS